Application for U.S. Patent:

FREQUENCY COEFFICIENT SCANNING PATHS FOR CODING DIGITAL VIDEO CONTENT

Inventors:

Limin Wang

and

Yue Yu

of

General Instrument Corporation

FREQUENCY COEFFICIENT SCANNING PATHS
FOR CODING DIGITAL VIDEO CONTENT

SPECIFICATION

CROSS-REFERENCE TO RELATED APPLICATION

[0001]   This application is related to U.S. Provisional Patent Application 60/416,139, filed on October 4, 2002, from which priority is claimed.

FIELD OF THE INVENTION

[0001]   The present invention relates to digital video encoding, decoding, and bitstream generation. More specifically, the present invention relates to scanning paths in transform-based coding as used in MPEG-4 Part 10 Advanced Video Coding/H.264, for example.

BACKGROUND OF THE INVENTION

[0002]   Video compression is used in many current and emerging products. It is at the heart of digital television set-top boxes ("STB"), digital satellite systems ("DSS"), high definition television ("HDTV") decoders, digital versatile disk ("DVD") players, video conferencing, Internet video and multimedia content, and other digital video applications. Without video compression, the number of bits required to represent digital video content can be extremely large, making it difficult or even impossible for the digital video content to be efficiently stored, transmitted, and/or viewed.

[0003]   The digital video content comprises a stream of pictures that can be displayed as an image on a television receiver, computer monitor, or other electronic device capable of displaying digital video content. A picture that is displayed in time before a particular picture

is in the "backward direction" in relation to the particular picture. Likewise, a picture that is displayed in time after a particular picture is in the "forward direction" in relation to the particular picture.

[0004]   Video compression is accomplished in a video encoding, or coding, process in which each picture is encoded as either a frame or as two fields. Each frame comprises a number of lines of spatial information. For example, a frame may contain 480 horizontal lines. Each field contains half the number of lines in the frame. For example, if the frame comprises 480 horizontal lines, each field comprises 240 horizontal lines. In a typical configuration, one of the fields comprises the odd numbered lines in the frame and the other field comprises the even numbered lines in the frame. The field that comprises the odd numbered lines will be referred to as the "top" field hereinafter and in the appended claims, unless otherwise specifically denoted. Likewise, the field that comprises the even numbered lines will be referred to as the "bottom" field hereinafter and in the appended claims, unless otherwise specifically denoted. The two fields can be interlaced together to form an interlaced frame.

[0005]   The general idea behind video coding is to remove data from the digital video content that is "non-essential." The decreased amount of data then requires less bandwidth for broadcast or transmission of the representation of the original signal.

[0006]   After the compressed (encoded) video data has been transmitted, it must be decoded, or decompressed. In this process, the transmitted video data is processed to generate approximation data that is substituted into the video data to replace the "non-essential" data that was removed in the above-mentioned coding process.

[0007]   Thus, video coding transforms the digital video content into a compressed form that can be stored using less space and transmitted using less bandwidth than uncompressed digital video content. It does so by taking advantage of temporal and spatial redundancies in the pictures of the video content. The resultant digital video content can then be stored in a storage medium such as a hard drive, DVD, or some other non-volatile storage unit.

[0008]   There are numerous video coding methods that compress the digital video content. Consequently, video coding standards have been developed to standardize the various video coding methods so that the compressed digital video content is rendered in formats that a majority of video encoders and decoders can recognize. For example, the Motion Picture Experts Group ("MPEG") and International Telecommunication Union ("ITU-T") have developed video coding standards that are in wide use today. Examples of these standards include the MPEG-1, MPEG-2, MPEG-4, ITU-T H.261, and ITU-T H.263 standards. However, with the current increased demand for higher resolutions, more complex graphical content, and faster transmission time, there exist a need for better video compression methods. To this end, a new video coding standard is currently being developed. This new video coding standard is called the MPEG-4 Part 10 Advanced Video Coding (AVC)/H.264 standard.

[0009]   Most modern video coding standards, including the MPEG-4 Part 10 AVC/H.264 standard, are based in part on a temporal prediction with a motion compensation ("MC") algorithm and a transform domain coding algorithm. Temporal prediction with motion compensation is used to remove temporal redundancy between successive pictures in a digital video broadcast. The temporal prediction with motion compensation algorithm typically utilizes one or two reference pictures to encode a particular picture. Thus, by comparing the

particular picture that is to be encoded with one of the reference pictures, the temporal

prediction with motion compensation algorithm can take advantage of the temporal

redundancy that exists between the reference picture and the particular picture that is to be

encoded, and encode the picture with a higher amount of compression than if the picture were

encoded without using the temporal prediction with motion compensation algorithm. One of

the reference pictures is in the backward direction in relation to the particular picture that is to

be encoded. The other reference picture is in the forward direction in relation to the particular

picture that is to be encoded.

[0010] Transform domain coding is used to remove spatial redundancy within each picture

or temporally predicted residual picture. A residual picture is the difference between a picture

and a picture that is temporally predicted from that picture. Each picture or temporally

predicted residual picture comprises a number of blocks of pixels. Each block refers to an N

by M group of pixels where N refers to the number of columns of pixels in the block and M

refers to the number of rows of pixels in the block. Each block in the picture or temporally

predicted residual picture is represented by an N by M array of luminance and chrominance

coefficients which correspond to each pixel in the blocks' N by M grid of pixels. Each

luminance coefficient represents the brightness level, or luminance, of its corresponding pixel.

Each block in the picture or temporally predicted residual picture is also represented by an N

by M array of chrominance coefficients which correspond to each pixel in the blocks' N by M

grid of pixels. Each chrominance coefficient represents the color content, or chrominance, of

its corresponding pixel. The term "picture" will be used hereinafter and in the appended

claims, unless otherwise specifically denoted, to mean either a picture or a temporally predicted residual picture.

[0011] Most pictures have smooth color variations, with the fine details being represented as sharp edges in between the smooth variations. The smooth variations in color can be termed as low frequency variations and the sharp variations as high frequency variations. The smooth variations in color, or low frequency components of the picture, constitute the base of an image, and the edges which give detail to the picture, or the high frequency components, add upon the smooth variations in color to refine the picture. The combination of the low and high frequency components results in a detailed image.

[0012] Typically, the values of the luminance coefficients only vary slightly between the most of the pixels in a particular picture. Consequentially, in many pictures, most pixels contain more of the low frequency component than the high frequency component. In other words, most of the energy of a signal containing the digital video content lies at low frequencies.

[0013] Transform domain coding takes advantage of the fact that most of the energy of a signal containing the digital video content lies at low frequencies. Transform domain coding transforms the luminance coefficients in each N by M array from the spatial domain to the frequency domain. The transformed N by M array comprises coefficients which represent energy levels in the frequency domain. As used hereinafter and in the appended claims, unless otherwise denoted, the coefficients of the transformed N by M array will be referred to as "frequency coefficients." Once the luminance coefficients have been transformed into frequency coefficients, various compression techniques can then be performed on the contents

of picture in the frequency domain that would otherwise be impossible to perform in the spatial domain.

[0014]   The N by M array of frequency coefficients is two dimensional and must be converted into a one dimensional array of frequency coefficients so that the encoder or decoder can use the frequency coefficients to encode or decode the picture. The encoder generates the one dimensional array of frequency coefficients by scanning the two dimensional array of frequency coefficients using a particular scanning path. The scanning path refers to the order in which the frequency coefficients in the two dimensional array are scanned and output by the encoder into the one dimensional array.

[0015]   A common scanning path that is used by an encoder to scan the frequency coefficients is a zig-zag scanning path. FIG. 1 illustrates two variations of zig-zag scanning paths that are currently used to scan a four by four array of frequency coefficients. As shown in FIG. 1, the first zig-zag scanning path 100 goes in a zig-zag order starting with an upper left coefficient (0) and ending with a lower right coefficient (15) of the array of frequency coefficients. The second zig-zag scanning path 101 is similar to the first in that it starts with the upper left coefficient (0) and ends with the lower right coefficient (15). However, as shown in FIG. 1, the two zig-zag scanning paths 100, 101 differ slightly in the order that the coefficients are scanned. FIG. 1 also shows one non-zig-zag scanning path 102 that is also prior art. Other prior art scanning paths for an 8 by 8 array of frequency coefficients can be found in MPEG-2 (Generic Coding of Moving Pictures and Associated Audio, Draft of International Standard, ISO/IEC 13818-2, March 1994).

[0016] It is preferable for the encoder to first scan the high-energy low frequency coefficients and then scan the low-energy high frequency coefficients. Scanning the low frequency coefficients before the high frequency coefficients places the low frequency coefficients before the high frequency coefficients in the resulting one dimensional array of coefficients. This particular order allows efficient coding and compression of the picture.

[0017] The zig-zag scanning path scans the two dimensional array of frequency coefficients without any significant bias towards the horizontal or vertical frequency directions. However, for interlaced material, energy tends to be concentrated along the vertical direction.

[0018] Thus, there is a need in the art for scanning paths that allow for more compression than do traditional zig-zag scanning paths. In particular, there is a need for a scanning path(s) which is biased in the vertical direction such that non-zero frequency coefficients would be captured first, thereby allowing for better compression efficiency.

SUMMARY OF THE INVENTION

[0019]   An objective of the present invention is to provide a system and method which scans frequency coefficients in a manner which is biased in the vertical direction, thereby allowing for better digital video compression efficiency.

[0020]   In order to achieve these objectives, as well as others which will become apparent in the disclosure below, in a first exemplary embodiment, the present invention provides a method of scanning frequency coefficients in a manner that is efficient for interlaced digital video content. The digital video content comprises a stream of pictures, slices, or macroblocks which can each be intra, predicted or bi-predicted pictures. The pictures, slices, or macroblocks comprise blocks of pixels in the configuration of a two dimensional array of two dimensional array frequency coefficients. The method of this exemplary embodiment comprises scanning each of the blocks of the two dimensional array of two dimensional array frequency coefficients in a manner that is vertically biased, thereby producing a one dimensional array of one dimensional array frequency coefficients. Further, the method of the present invention may alternatively scan one dimensional array frequency coefficients of a one dimensional array, coded in the vertically biased manner of the present invention, thereby producing a two dimensional array of two dimensional array frequency coefficients.

[0021]   In a second exemplary embodiment, the present invention provides an encoder that scans frequency coefficients in a manner that is efficient for interlaced digital video content. The digital video content comprises a stream of pictures, slices, or macroblocks which can each be intra, predicted or bi-predicted pictures. The pictures, slices, or macroblocks comprise blocks of pixels in the configuration of a two dimensional array of two dimensional array

frequency coefficients. The encoder scans each of the blocks of the two dimensional array of two dimensional array frequency coefficients in a manner that is vertically biased, thereby producing a one dimensional array of one dimensional array frequency coefficients.

[0022] In yet a third exemplary embodiment, the present invention provides a decoder that scans a one dimensional array of one dimensional array frequency coefficients and produces a two dimensional array of two dimensional array frequency coefficients in a manner that is efficient for interlaced digital video content. In this exemplary embodiment, the digital video content comprises a stream of pictures, slices, or macroblocks which can each be intra, predicted or bi-predicted pictures. The pictures, slices, or macroblocks comprise blocks of pixels in the configuration of a one dimensional array of one dimensional array frequency coefficients. The decoder scans one dimensional array frequency coefficients of a one dimensional array, coded in the vertically biased manner of the present invention, thereby producing a two dimensional array of two dimensional array frequency coefficients.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023]   For a complete understanding of the present invention and the advantages thereof, reference is now made to the following description taken in conjunction with the accompanying drawings in which like reference numbers indicate like features, components and method steps, and wherein:

[0024]   FIG. 1 is prior art and illustrates two variations of zig-zag scanning paths and a non-zig-zag scanning path that are currently used to scan a four by four array of frequency coefficients;

[0025]   FIG. 2 illustrates an exemplary sequence of three types of pictures according to an embodiment of the present invention, as defined by an exemplary video coding standard such as the MPEG-4 Part 10 AVC/H.264 standard;

[0026]   FIG. 3 illustrates that each picture is preferably divided into one or more slices consisting of macroblocks;

[0027]   FIG. 4 illustrates that a macroblock can be further divided into smaller sized blocks;

[0028]   FIG. 5 illustrates a preferable method of transform domain coding in accordance with an exemplary embodiment of the present invention;

[0029]   FIG. 6 illustrates a preferable scanning path for a four by four pixel block's frequency coefficient array in accordance with an exemplary embodiment of the present invention;

[0030]   FIG. 7 illustrates a preferable scanning path for a four by eight pixel block's frequency coefficient array in accordance with an exemplary embodiment of the present invention;

[0031] FIG. 8 illustrates a preferable scanning path for an eight by four pixel block's frequency coefficient array in accordance with an exemplary embodiment of the present invention.

[0032] FIG. 9 illustrates a preferable scanning path for an eight by eight pixel block's frequency coefficient array in accordance with an exemplary embodiment of the present invention.

[0033] Throughout the drawings, identical reference numbers designate similar, but not necessarily identical, elements.

DESCRIPTION OF A PRESENTLY PREFERRED EMBODIMENT

[0034]   The present invention provides methods for scanning frequency coefficients from a

two dimensional array of two dimensional frequency coefficients to produce a one

dimensional array of one dimensional array frequency coefficients ("encoding scan"). The

present invention also provides methods for scanning/assigning frequency coefficients from a

one dimensional array of one dimensional frequency coefficients to produce a two

dimensional array of two dimensional array frequency coefficients; the mirror image

("decoding scan"). Further, the present invention provides for an encoder and a decoder,

featuring the encoding and decoding scans of the present invention, respectively. In addition,

the present invention provides for systems containing both at least one encoder and at least

one decoder which employ the encoding and decoding scans, respectively, e.g., transmission

systems, transcoders, etc.

[0035]   These methods can be used in any digital video coding algorithm. In particular, they

can be implemented in the MPEG-4 Part 10 AVC/H.264 video coding standard.

[0036]   As noted above, the MPEG-4 Part 10 AVC/H.264 standard is a new standard for

encoding and compressing digital video content. The documents establishing the MPEG-4

Part 10 AVC/H.264 standard are hereby incorporated by reference, including the "Joint Final

Committee Draft ("JFCD") of Joint Video Specification" issued on August 10, 2002 by the

Joint Video Team ("JVT"). (ITU-T Rec. H.264 & ISO/IEC 14496-10 AVC). The JVT

consists of experts from MPEG and ITU-T. Due to the public nature of the MPEG-4 Part 10

AVC/H.264 standard, the present specification will not attempt to document all the existing

aspects of MPEG-4 Part 10 AVC/H.264 video coding, relying instead on the incorporated

specifications of the standard.

[0037]   The systems, devices and methods of the present invention can be used in any

general digital video coding algorithm or system requiring coefficient scanning. Further, the

methods of the present invention can be modified and used to handle the extraction of

frequency coefficients from a two dimensional array of two dimensional array frequency

coefficients as best serves a particular standard or application.

[0038]   Using the drawings, the following exemplary embodiments of the present invention

will now be explained.

[0039]   As shown in FIG. 2, there are preferably three types of pictures that can be used in

the video coding method. Three types of pictures are defined to support random access to

stored digital video content while exploring the maximum redundancy reduction using

temporal prediction with motion compensation. The three types of pictures are intra ("I")

pictures 200, predicted ("P") pictures 202a-b, and bi-predicted ("B") pictures 201a-d. An I

picture 200 provides an access point for random access to stored digital video content. Intra

pictures 200 are encoded without referring to reference pictures and can be encoded with

moderate compression.

[0040]   A P picture 202a-b is encoded using an I, P, or B picture that has already been

encoded as a reference picture. The reference picture can be in either the forward or backward

temporal direction in relation to the P picture that is being encoded.   The predicted pictures

202a-b can be encoded with more compression than the I pictures 200.

[0041] A B picture 201a-d is encoded using two temporal reference pictures. In accordance with an exemplary embodiment of the present invention, the two temporal reference pictures can be in the same or different temporal direction in relation to the B picture that is being encoded. B pictures 201a-d can be encoded with the most compression out of the three picture types.

[0042] Reference relationships 203 between the three picture types are illustrated in FIG. 2. For example, the P picture 202a can be encoded using the encoded I picture 200 as its reference picture. The B pictures 201a-d can be encoded using the encoded I picture 200 and the encoded P pictures 202a-b as its reference pictures, as shown in FIG. 2. In accordance with an exemplary embodiment of the present invention, encoded B pictures 201a-d can also be used as reference pictures for other B pictures that are to be encoded. For example, the B picture 201c of FIG. 2 is shown with two other B pictures 201b and 201d as its reference pictures.

[0043] The number and particular order of the I 200, B 201a-d, and P 202a-b pictures shown in FIG. 2 are given as an exemplary configuration of pictures, but are not necessary to implement the present invention. Any number of I, B, and P pictures can be used in any order to best serve a particular application. The MPEG-4 Part 10 AVC/H.264 standard does not impose any limit to the number of B pictures between two reference pictures nor does it limit the number of pictures between two I pictures.

[0044] FIG. 3 shows that each picture 300 is preferably divided into slices consisting of macroblocks. A slice 301 is a group of macroblocks and a macroblock 302 is a rectangular group of pixels. As shown in FIG. 3, a preferable macroblock 302 size is 16 by 16 pixels.

[0045]   Each interlaced picture, slice, or macroblock in a stream of pictures that is to be encoded can be encoded using adaptive frame/field ("AFF") coding. In AFF coding, each picture, slice, or macroblock in a stream of pictures that is to be encoded is encoded in either frame mode or in field mode, regardless of the encoding mode of the previous picture, slice, or macroblock. If a picture, slice, or macroblock is encoded in frame mode, the two fields that make up an interlaced frame are coded jointly. Conversely, if a picture, slice, or macroblock is encoded in field mode, the two fields that make up an interlaced frame are coded separately. The encoder determines which type of coding, frame mode coding or field mode coding, is more advantageous for each picture, slice, or macroblock and chooses that type of encoding for the picture, slice, or macroblock. The exact method of choosing between frame mode and field mode is not critical to the present invention and will not be detailed herein.

[0046]   FIG. 4 shows that a macroblock can be further divided into smaller sized blocks. For example, as shown in FIG. 4, a macroblock can be further divided into block sizes of 16 by 8 pixels 400, 8 by 16 pixels 401, or 8 by 8 pixels 402. A block size of 8 by 8 pixels 402 can be further subdivided into block sizes of 8 by 4 pixels 403, 4 by 8 pixels 404, or 4 by 4 pixels 405.

[0047]   A picture that is to be encoded using transform domain coding can sometimes be encoded with better picture quality or more compression efficiency if the transform domain coding is performed on the smaller block sizes of FIG. 4 rather than on the macroblock itself. Some digital video coding algorithms allow for variable block size transforms. Variable block size transform coding means that the transform domain coding can be performed on blocks of varying sizes. For example, transform domain coding can be performed on 4 by 4 pixel

blocks 405 for a particular macroblock and on 4 by 8 pixel blocks 404 for a different

macroblock. Transform domain coding on the following block sizes can be implemented in

accordance with an exemplary embodiment of the present invention: 4 by 4 pixels 405, 8 by 4

pixels 403, 4 by 8 pixels 404, and 8 by 8 pixels 402.

[0048]   FIG. 5 illustrates a preferable method of transform domain coding in accordance

with an exemplary embodiment of the present invention. As shown in FIG. 5, a transform 500

is performed on a block's N by M array of luminance or chrominance coefficients. The N by

M array of luminance or chrominance coefficients comprises the coefficients that represent the

luminance or chrominance of the pixels in the N by M block. The N by M array of luminance

or chrominance coefficients can be a 4 by 4 array, 4 by 8 array, 8 by 4 array, or an 8 by 8 array

in accordance with this exemplary embodiment of the present invention.

[0049]   The discrete cosine transform ("DCT") is an example of a transform and is similar to

the discrete Fourier transform. The DCT transforms the N by M array of luminance or

chrominance coefficients from the spatial domain to the frequency domain. The general

equation for a two dimensional, N by M, DCT can be defined by the following equation:

[0050] $$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}}\left(\frac{2}{M}\right)^{\frac{1}{2}}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\Lambda(i).\Lambda(j).\cos\left[\frac{\pi.u}{2.N}(2i+1)\right]\cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

where $\Lambda(\xi) = \begin{cases} \dfrac{1}{\sqrt{2}} & for \xi = 0 \\ 1 & otherwise \end{cases}$.

[0051]   In the above equations, $f(i,j)$ represents the luminance or chrominance value of the

pixel in column i and row j of the N by M array of luminance coefficients. $F(u,v)$ is the

corresponding frequency coefficient in column u and row v in the N by M array of frequency

coefficients. For most images, much of the signal energy lies at low frequencies. In general, the low frequency coefficients appear in the upper left corner of the N by M array of frequency coefficients. The high frequency coefficients usually appear in the lower right corner of the N by M array of frequency coefficients.

[0052] After the luminance or chrominance coefficients have been converted to frequency coefficients by the transform 500, the frequency coefficients are quantized 501, as shown in FIG. 5. Quantization 501 is performed on the frequency coefficients so that the number of bits that must be encoded is reduced. This allows for more compression.

[0053] One example of the quantization process 501 consists of dividing each $F(u,v)$ by a constant, q(u,v). A table of q(u,v) is called a quantization table. An exemplary, but not exclusive, quantization table for an 8 by 8 array of frequency coefficients is shown in Table 1 below:

Table 1: Eight by eight quantization table

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	36	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

[0054] Similar quantization tables can be constructed for the other sizes of the N by M frequency coefficient array. As shown in exemplary quantization table, the constants that divide each $F(u,v)$ are larger in value in the lower right corner of the quantization table than

they are in the upper left corner. An important result of the quantization process is that many of the high frequency coefficients are quantized to a value of zero.

[0055] Returning to FIG. 5, the quantized frequency coefficients are scanned 502 by the encoder to convert them from a two dimensional array of quantized frequency coefficients to a one dimensional array of quantized frequency coefficients. Preferable scanning paths will be described in more detail in connection with FIGs. 6-9 below.

[0056] After the quantized frequency coefficients have been scanned into the one dimensional array, they can be encoded 503, as shown in FIG. 5. An exemplary encoding 503 process preferably encodes the quantized frequency coefficients in the one dimensional array into a sequence of run-level pairs. The run is defined as the distance between two non-zero quantized frequency coefficients in the one dimensional array. The level is the non-zero value immediately following a sequence of zeros. This type of coding produces a compact representation of the quantized frequency coefficients because a large number of the quantized coefficients have a value of zero. The run-level pairs can be further compressed using entropy coding. One method of entropy coding is described in detail in the MPEG-4 Part 10 AVC/H.264 standard. MPEG-4 Part 10 AVC/H.264 also uses context-adaptive binary arithmetic coding ("CABAC").

[0057] Preferable scanning paths for scanning the frequency coefficients in the two dimensional array into a one dimensional array of frequency coefficients (encoding scan) will now be explained in connection with FIGs. 6-9. An ideal scanning path in any block size would group all the non-zero quantized frequency coefficients together in the one dimensional array followed by all the quantized frequency coefficients that have values of zero. However,

in practice, a preferable scanning path can only group together a majority of non-zero quantized frequency coefficients. For interlaced material, the non-zero quantized frequency coefficients tend to be concentrated along the vertical direction and a vertically biased scanning path may be preferable.

[0058]   FIGs. 6-9 show preferable scanning path orders for a 4 by 4 pixel block, 4 by 8 pixel block, 8 by 4 pixel block, and an 8 by 8 pixel block, respectively. In the following descriptions, $n=0,1,\ldots,N-1$, where n is a variable that represents the pixel column number in the block as well as the corresponding frequency coefficient column number in the corresponding frequency coefficient array. N is the total number of pixel columns in the block and the total number of frequency coefficient columns in the frequency coefficient array. The left-most column number is 0 and the right-most column number is N-1. Likewise, $m=0,1,\ldots,M-1$, where m is a variable that represents the pixel row number in the block as well as the corresponding frequency coefficient row number in the corresponding frequency coefficient row number in the corresponding frequency coefficient array. M is the total number of pixel rows in the block and the total number of frequency coefficient rows in the frequency coefficient array. The top row number is 0 and the bottom row number is M-1. The scanning paths of FIGs. 6-9 are skewed, or biased, in the vertical direction and result in more compression than traditional zig-zag scanning paths in many applications, including interlaced video encoding.

[0059]   FIG. 6 shows a preferable scanning path for a 4 by 4 pixel block's frequency coefficient array, where N=4 and M=4. The numbers in FIG. 6 represent the frequency coefficient scanning order. For example, the frequency coefficient corresponding to the top

left pixel is the first frequency coefficient to get scanned and is thus labeled with a 0. The

frequency coefficient corresponding to the bottom right pixel is the last frequency coefficient

to get scanned and is thus labeled with a 15. Table 2 lists the frequency coefficient scanning

order and the corresponding values for n and m.

Table 2: Four by four pixel block scanning order

Frequency Coefficient Scanning Order	n	m
0	0	0
1	0	1
2	1	0
3	0	2
4	0	3
5	1	1
6	1	2
7	1	3
8	2	0
9	2	1
10	2	2
11	2	3
12	3	0
13	3	1
14	3	2
15	3	3

[0060] FIG. 7 shows a preferable scanning path for a 4 by 8 pixel block's frequency

coefficient array, where N=4 and M=8. The numbers in FIG. 7 represent the frequency

coefficient scanning order. For example, the frequency coefficient corresponding to the top

left pixel is the first frequency coefficient to get scanned and is thus labeled with a 0. The

frequency coefficient corresponding to the bottom right pixel is the last frequency coefficient

to get scanned and is thus labeled with a 31. Table 3 lists the frequency coefficient scanning

order and the corresponding values for n and m.

Table 3: Four by eight pixel block scanning order

Frequency Coefficient Scanning Order	n	m
0	0	0
1	0	1
2	0	2
3	0	3
4	1	0
5	1	1
6	1	2
7	0	4
8	0	5
9	0	6
10	0	7
11	1	3
12	2	0
13	2	1
14	2	2
15	1	4
16	1	5
17	1	6
18	1	7
19	2	3
20	3	0
21	3	1
22	3	2
23	2	4
24	2	5
25	2	6
26	2	7
27	3	3
28	3	4
29	3	5
30	3	6
31	3	7

[0061]   FIG. 8 shows a preferable scanning path for an 8 by 4 pixel block's frequency coefficient array, where N=8 and M=4. The numbers in FIG. 8 represent the frequency coefficient scanning order. For example, the frequency coefficient corresponding to the top left pixel is the first frequency coefficient to get scanned and is thus labeled with a 0. The frequency coefficient corresponding to the bottom right pixel is the last frequency coefficient to get scanned and is thus labeled with a 31. Table 4 lists the frequency coefficient scanning order and the corresponding values for n and m.

Table 4: Eight by four pixel block scanning order

Frequency Coefficient Scanning Order	n	m
0	0	0
1	0	1
2	1	0
3	0	2
4	0	3
5	1	1
6	2	0
7	1	2
8	1	3
9	2	1
10	3	0
11	2	2
12	2	3
13	3	1
14	4	0
15	3	2
16	3	3
17	4	1
18	5	0
19	4	2
20	4	3
21	5	1
22	6	0

23	5	2
24	5	3
25	6	1
26	7	0
27	6	2
28	6	3
29	7	1
30	7	2
31	7	3

[0062]   FIG. 9 shows a preferable scanning path for an 8 by 8 pixel block's frequency coefficient array, where N=8 and M=8. The numbers in FIG. 9 represent the frequency coefficient scanning order. For example, the frequency coefficient corresponding to the top left pixel is the first frequency coefficient to get scanned and is thus labeled with a 0. The frequency coefficient corresponding to the bottom right pixel is the last frequency coefficient to get scanned and is thus labeled with a 63. Table 5 lists the frequency coefficient scanning order and the corresponding values for n and m.

Table 5: Eight by eight pixel block scanning order

Frequency Coefficient Scanning Order	n	m
0	0	0
1	0	1
2	0	2
3	1	0
4	1	1
5	0	3
6	0	4
7	1	2
8	2	0
9	1	3
10	0	5
11	0	6

24

12	0	7
13	1	4
14	2	1
15	3	0
16	2	2
17	1	5
18	1	6
19	1	7
20	2	3
21	3	1
22	4	0
23	3	2
24	2	4
25	2	5
26	2	6
27	2	7
28	3	3
29	4	1
30	5	0
31	4	2
32	3	4
33	3	5
34	3	6
35	3	7
36	4	3
37	5	1
38	6	0
39	5	2
40	4	4
41	4	5
42	4	6
43	4	7
44	5	3
45	6	1
46	6	2
47	5	4
48	5	5
49	5	6
50	5	7
51	6	3
52	7	0
53	7	1

54	6	4
55	6	5
56	6	6
57	6	7
58	7	2
59	7	3
60	7	4
61	7	5
62	7	6
63	7	7

[0063]   The above-described scanning paths relate to methods of encoding scans (see

definition above) in accordance with the present invention. Such scanning paths may be

implemented in an encoder in accordance with the present invention.

[0064]   As illustrated above, the 4 by 4, 4 by 8, 8 by 4, and 8 by 8 pixel block encoding

scans produce a one dimensional array of 16, 32, 32, or 64 one dimensional array frequency

coefficients, respectively. Similarly, the scanning/assignment paths for decoding (method of

decoding scans) for each one dimensional array of 16, 32, 32, and 64 frequency coefficients is

the mirror image of the encoding scanning paths for the 4 by 4, 4 by 8, 8 by 4, and 8 by 8 pixel

blocks, respectively. For the preferred decoding scan embodiments described below, one

dimensional array frequency coefficients are denoted by the variable "P".

[0065]   In particular, in accordance with the present invention a preferred scanning

path/assignment for a one dimensional array of 16 frequency coefficients, where a 4 by 4 two

dimensional array of frequency coefficients is desired, is effectuated by assigning two

dimensional frequency coefficient values (N, M) for each P value (frequency coefficient) in

the one dimensional in the numerical sequential order of P value, as illustrated below in Table

6.

CONFIDENTIAL
ATTORNEY-CLIENT PRIVILEGED

Table 6: 1D of 16 P to four by four pixel block scanning/assignment order

p	n	m
0	0	0
1	0	1
2	1	0
3	0	2
4	0	3
5	1	1
6	1	2
7	1	3
8	2	0
9	2	1
10	2	2
11	2	3
12	3	0
13	3	1
14	3	2
15	3	3

[0066]   In accordance with the present invention a preferred scanning path/assignment for a one dimensional array of 32 frequency coefficients, where a 4 by 8 two dimensional array of frequency coefficients is desired, is effectuated by assigning two dimensional frequency coefficient values (N, M) for each P value (frequency coefficient) in the one dimensional array in the numerical sequential order of P value, as illustrated below in Table 7.

Table 7: 1D of 32 P to four by eight pixel block scanning/assignment order

P	n	m
0	0	0
1	0	1
2	0	2
3	0	3
4	1	0
5	1	1
6	1	2

7	0	4
8	0	5
9	0	6
10	0	7
11	1	3
12	2	0
13	2	1
14	2	2
15	1	4
16	1	5
17	1	6
18	1	7
19	2	3
20	3	0
21	3	1
22	3	2
23	2	4
24	2	5
25	2	6
26	2	7
27	3	3
28	3	4
29	3	5
30	3	6
31	3	7

[0067]   In accordance with the present invention a preferred scanning path/assignment for a

one dimensional array of 32 frequency coefficients, where a 8 by 4 two dimensional array of

frequency coefficients is desired, is effectuated by assigning two dimensional frequency

coefficient values (N, M) for each P value (frequency coefficient) in the one dimensional array

in the numerical sequential order of P value, as illustrated below in Table 8.

Table 8: 1D of 32 P to eight by four pixel block scanning/assignment order

p	n	m
0	0	0
1	0	1
2	1	0
3	0	2
4	0	3
5	1	1
6	2	0
7	1	2
8	1	3
9	2	1
10	3	0
11	2	2
12	2	3
13	3	1
14	4	0
15	3	2
16	3	3
17	4	1
18	5	0
19	4	2
20	4	3
21	5	1
22	6	0
23	5	2
24	5	3
25	6	1
26	7	0
27	6	2
28	6	3
29	7	1
30	7	2
31	7	3

[0068]   In accordance with the present invention a preferred scanning path/assignment for a

one dimensional array of 64 frequency coefficients, where a 8 by 8 two dimensional array of

frequency coefficients is desired, is effectuated by assigning two dimensional frequency

coefficient values (N, M) for each P value (frequency coefficient) in the one dimensional array

in the numerical sequential order of P value, as illustrated below in Table 9.

Table 9: 1D of 32 P to four by eight pixel block scanning/assignment order

p	n	m
0	0	0
1	0	1
2	0	2
3	1	0
4	1	1
5	0	3
6	0	4
7	1	2
8	2	0
9	1	3
10	0	5
11	0	6
12	0	7
13	1	4
14	2	1
15	3	0
16	2	2
17	1	5
18	1	6
19	1	7
20	2	3
21	3	1
22	4	0
23	3	2
24	2	4
25	2	5
26	2	6
27	2	7
28	3	3
29	4	1
30	5	0
31	4	2
32	3	4
33	3	5

34	3	6
35	3	7
36	4	3
37	5	1
38	6	0
39	5	2
40	4	4
41	4	5
42	4	6
43	4	7
44	5	3
45	6	1
46	6	2
47	5	4
48	5	5
49	5	6
50	5	7
51	6	3
52	7	0
53	7	1
54	6	4
55	6	5
56	6	6
57	6	7
58	7	2
59	7	3
60	7	4
61	7	5
62	7	6
63	7	7

[0069]   The above systems and methods may be implemented by many computer languages commonly known in the art and may operate on many computer platforms which include both volatile and non-volatile memory storage devices.

[0070]   Although the invention has been described herein by reference to an exemplary embodiment thereof, it will be understood that such embodiment is susceptible of modification

32

and variation without departing from the inventive concepts disclosed. All such modifications

and variations, therefore, are intended to be encompassed within the spirit and scope of the

appended claims.